Conformational Temperature-Dependent Behavior of a Histone H2AX: A Coarse-Grained Monte Carlo Approach Via Knowledge-Based Interaction Potentials

Miriam Fritsche[1]*, Ras B. Pandey[2], Barry L. Farmer[3], Dieter W. Heermann[1,4,5]

1 Institute for Theoretical Physics, University of Heidelberg, Heidelberg, Germany, 2 Department of Physics and Astronomy, University of Southern Mississippi, Hattiesburg, Mississippi, United States of America, 3 Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright Patterson Air Force Base, Ohio, United States of America, 4 Institute for Molecular Biophysics, The Jackson Laboratory, Bar Harbor, Maine, United States of America, 5 Interdisciplinary Center for Scientific Computing, Heidelberg, Germany

Abstract

Histone proteins are not only important due to their vital role in cellular processes such as DNA compaction, replication and repair but also show intriguing structural properties that might be exploited for bioengineering purposes such as the development of nano-materials. Based on their biological and technological implications, it is interesting to investigate the structural properties of proteins as a function of temperature. In this work, we study the spatial response dynamics of the histone H2AX, consisting of 143 residues, by a coarse-grained bond fluctuating model for a broad range of normalized temperatures. A knowledge-based interaction matrix is used as input for the residue-residue Lennard-Jones potential.

We find a variety of equilibrium structures including global globular configurations at low normalized temperature ($T^* = 0.014$), combination of segmental globules and elongated chains ($T^* = 0.016, 0.017$), predominantly elongated chains ($T^* = 0.019, 0.020$), as well as universal SAW conformations at high normalized temperature ($T^* \geq 0.023$). The radius of gyration of the protein exhibits a non-monotonic temperature dependence with a maximum at a characteristic temperature ($T_c^* = 0.019$) where a crossover occurs from a positive (stretching at $T^* \leq T_c^*$) to negative (contraction at $T^* \geq T_c^*$) thermal response on increasing T^*.

Introduction

Our perspective of proteins is more and more changing from static to dynamical entities [1]. Increasing efforts are undertaken to better understand the functioning of biological macromolecules with the aim of engineering them for technological applications [1]. In fact, proteins are powerful therapeutic agents [2,3] and recent developments have stressed their impact on the bioengineering of nanomaterials, especially in biomedical imaging, drug delivery, biosensing and the design of functional nanocomposites [4,5]. To this end, protein self-assembly in suitable media offers unique advantages in the fabrication of protein-based nanodevices and avoids cost-intensive manufacturing processes [1]. However, thermally stable proteins are essential for bionanotechnological applications [6–8]. Thus, based on the biological [9–12] and technological implications [1,13,14], it is interesting to investigate the structural properties of proteins as a function of temperature [15,16].

Due to the unique sequence of residues a complex and specific structural response is expected. Consequently, the study of the structure of proteins has been of interest for many years, particularly exploiting computer simulations via idealized models [6,7,17–28]. However, despite considerable effort, a complete understanding of the fundamental issue of how proteins fold to its native structure still remains an open question from a first principle approach [6,7,17–28].

In this work, we explore how the protein H2AX, a variant of the histone H2A, conforms and responds to temperature changes. The protein H2AX plays a vital role in cellular functioning: Its serine 139 phosphorylated form, γ-H2AX [29], is one of the earliest repair responses to DNA double-strand breaks which can lead to mutations that in turn are a cause of cancer and hereditary disease [30]. Notably, it was found that γ-H2AX increases in a variety of conditions in relation to double-strand break generation processes, including radiation and high temperature [31]. Moreover, a recent work has shown that postranslational modifications in histones underlie heat acclimation-mediated cytoprotective memory [32].

Regarding the above mentioned technological implications of proteins in general and the biological impact of the histone H2AX in particular the study of their temperature-dependent structural changes is worthwhile. Thus, we consider a coarse-grained model [18,20] of this protein chain where knowledge-based residue-residue interactions [19,33] are employed and effects of temperature are explored.

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **19 MAR 2012**	2. REPORT TYPE	3. DATES COVERED **00-00-2012 to 00-00-2012**
4. TITLE AND SUBTITLE **Conformational Temperature-Dependent Behavior of a Histone H2AX: A Coarse-Grained Monte Carlo Approach Via Knowledge-Based Interaction Potentials**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory,Materials and Manufacturing Directorate,Wright Patterson AFB,OH,45433**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		

14. ABSTRACT

Histone proteins are not only important due to their vital role in cellular processes such as DNA compaction, replication and repair but also show intriguing structural properties that might be exploited for bioengineering purposes such as the development of nano-materials. Based on their biological and technological implications, it is interesting to investigate the structural properties of proteins as a function of temperature. In this work, we study the spatial response dynamics of the histone H2AX, consisting of 143 residues, by a coarse-grained bond fluctuating model for a broad range of normalized temperatures. A knowledge-based interaction matrix is used as input for the residue-residue Lennard-Jones potential. We find a variety of equilibrium structures including global globular configurations at low normalized temperature (T ~0:014), combination of segmental globules and elongated chains (T ~0:016,0:017), predominantly elongated chains (T ~0:019,0:020), as well as universal SAW conformations at high normalized temperature (T ?0:023). The radius of gyration of the protein exhibits a non-monotonic temperature dependence with a maximum at a characteristic temperature (T c~0:019) where a crossover occurs from a positive (stretching at T ?T c) to negative (contraction at T ?T c) thermal response on increasing T .

15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

Analysis

The high dimensional space of protein conformations as well as the complexity of the energy surface make coarse-graining almost unavoidable in modeling the global structure and dynamics of proteins no matter whether one choses an all-atom approach, a minimalist description or a combination thereof [22–28].

In this work, we present a computer simulation study of the protein H2AX applying the bond-fluctuation method [34], which has been applied successfully to model the static and dynamical properties of polymer systems in several investigations [35]. It is a lattice algorithm with the advantage of avoiding non-ergodicity and its computational efficiency renders it more attractive than off-lattice models. The polymer evolves on a cubic lattice, where each monomer blocks eight lattice sites, which then cannot be occupied by other monomers. Monomers are connected by fluctuating bond vectors of lengths 2, $\sqrt{5}$, $\sqrt{6}$, 3 and $\sqrt{10}$, leading to an average bond length of $\langle b \rangle = 2.7$. A Monte Carlo move consists of randomly choosing one monomer to be moved to a randomly chosen lattice direction by one lattice unit. This trial motion is only accepted if neither the excluded volume constraint nor the restriction on the allowed range of bond lengths is violated. By additionally considering the Metropolis transition probability [35] for accepting or rejecting a move we include effects due to the finite interaction energies explained in the following [34]. Thus, the simulation method produces unbiased results, takes into account excluded volume interactions and ensures that no bond crossings can occur.

The H2AX protein consists of 143 residues shown in Table 1. Each residue is described by a monomer of the bond-fluctuating protein chain [18,20]. This is a simplified representation of a residue without the all-atom structural details but the specificity of each residue is captured via the applied residue-residue interactions [21]. Moreover, our approach has the advantage of computational efficiency allowing for the covering large (biological) scales. In fact, the bond-fluctuation method has recently been used to study the conformational relaxation into native structure of a general HP protein chain [36] and even a specific protein, sensory rhodopsin, without severe constraints [37].

Apart from excluded volume interactions, each residue interacts with the neighboring residues within the range r_c using a generalized Lennard-Jones (LJ) potential

$$U_{i,j} = \left[|\varepsilon_{ij}| \left(\frac{\sigma}{r_{ij}} \right)^{12} - \varepsilon_{ij} \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \quad \text{for } r_{ij} < r_c, \quad (1)$$

where r_{ij} is the distance between the residues i and j and $r_c = \sqrt{8}$ and $\sigma = 1$ in units of the lattice constant. The strength of the pair potential ε_{ij} is unique for each pair of residues with appropriate positive (repulsive) or negative (attractive) values (for more detail on the force field see [19,38]). In contrast to our recent study of a HIV protease with a coarse-grained approach involving the relative hydropathy index of each amino acid as well as results from all-atom simulations [18], in this work, we use a knowledge-based interaction matrix for the residue-residue pair interactions. The knowledge-based interaction potential matrix is derived from an ensemble of a large number of protein structures in the protein data bank (PDB). A number of such interaction tables are frequently used to investigate a range of questions related to protein structure including protein folding which has been studied extensively with a variety of models and methods involving all-atom details to minimalist coarse-grained descriptions [22–28,39–44]. We resort here to the classic residue-residue contact interaction table [19] which is employed in studying scaffolding of short peptides [20].

Even though the knowledge-based matrix elements ε_{ij} are simplified estimates derived from residue-residue contacts we are confident that the phenomenological interaction matrix implicitly takes into account the secondary and tertiary structure of the proteins [45,46]. In fact, several factors need to be considered for the crystallization of a protein sample subject to an X-ray christallograpy study, which include protein purity, pH, concentration of the protein, the temperature, and precipitants. X-ray crystallographic images of several thousands of proteins from the PDB are used to derive the residue-residue interaction matrix which is applied as an input for our potential. In such a huge ensemble of proteins, residues in secondary and tertiary structures are expected be well represented by effective residue-residue interactions at various temperatures when applying a coarse-grained protein model.

Each randomly selected residue performs its stochastic movement according to the Metropolis algorithm subject to excluded volume constraints and the limits on changes in the covalent (i.e. peptide) bond length as in our previous studies [18,20]. A

Table 1. Sequence of residues of the histone H2AX.

[1] M	[2] S	[3] G	[4] R	[5] G	[6] K	[7] T	[8] G	[9] G	[10] K	[11] A	[12] R
[13] A	[14] K	[15] A	[16] K	[17] S	[18] R	[19] S	[20] S	[21] R	[22] A	[23] G	[24] I
[25] Q	[26] F	[27] P	[28] V	[29] G	[30] R	[31] V	[32] H	[33] R	[34] L	[35] L	[36] R
[37] K	[38] G	[39] H	[40] Y	[41] A	[42] E	[43] R	[44] V	[45] G	[46] A	[47] G	[48] A
[49] P	[50] V	[51] Y	[52] L	[53] A	[54] A	[55] V	[56] L	[57] E	[58] Y	[59] L	[60] T
[61] A	[62] E	[63] I	[64] L	[65] E	[66] L	[67] A	[68] G	[69] N	[70] A	[71] A	[72] R
[73] D	[74] N	[75] K	[76] K	[77] T	[78] R	[79] I	[80] I	[81] P	[82] R	[83] H	[84] L
[85] Q	[86] L	[87] A	[88] I	[89] R	[90] N	[91] D	[92] E	[93] E	[94] L	[95] N	[96] K
[97] L	[98] L	[99] G	[100] G	[101] V	[102] T	[103] I	[104] A	[105] Q	[106] G	[107] G	[108] V
[109] L	[110] P	[111] N	[112] I	[113] Q	[114] A	[115] V	[116] L	[117] L	[118] P	[119] K	[120] K
[121] T	[122] S	[123] A	[124] T	[125] V	[126] G	[127] P	[128] K	[129] A	[130] P	[131] S	[132] G
[133] G	[134] K	[135] K	[136] A	[137] T	[138] Q	[139] A	[140] S	[141] Q	[142] E	[143] Y	

Hydrophobic residues are pink, polar residues are gold and electrostatic ones are blue.

randomly selected residue at a site i is moved to one of its randomly selected neighboring lattice sites j with the Boltzmann probability $\exp(-\Delta E_{ij}/T^*)$, where $\Delta E_{ij} = E_j - E_i$ is the change in energy between the attempted E_j and current E_i configuration [35].

T^* is the normalized temperature in units of the Boltzmann constant k_B and the energy ε_{ij}. Due to the lack of calibration with experimental data it is not possible to quantitatively relate the temperature T^* to physical temperature values. However, since the interresidue contact energies ε_{ij} allow for the calculation of realistic conformational energies of amino acids sequences in a number of different folds [38] we are able to relate temperature changes qualitatively to changes in the structural properties of proteins. Thus, our coarse-grained protein model provides a (qualitative) framework for understanding the temperature-dependent response of proteins which can so far not be gained by experimental testing.

Initially, the protein chain is placed in a random conformation with excluded volume constraints. Simulations are then performed for a sufficiently long time (typically 10^7 time steps) with 150 independent samples. While one can monitor (thermodynamic) quantities (such as the radius of gyration or the energy) in a simulation in order to make sure that the system has reached asymptotic steady state, one has to take into account that the protein may or may not be in equilibrium due to the possibility of metastability (caused by frustration). Gerstman and Chapagain [47] provide an estimate for the time which is required for a protein to undergo the transition from a random coil to its native state. Using a simplified coarse-grained model and introducing a propensity energy to constrain appropriate segmental structures they suggest that 10^7 time steps (corresponding to about 0.01 sec) is large enough for a protein to reach its native structure. In our simulation, the protein chain is initially in a random coil configuration (with excluded volume constraints) and it takes about 10^7 time steps to reach an equilibrium conformation. Thus, while one has to take into account that this approximation might fail when considering additional details such as an effective medium etc., the time scale by Gerstman and Chapagain could be a rough estimate of the order of magnitude for our simulation. Different lattice sizes are used to test for finite size effects. Most of the data presented here are generated on a lattice of size $64 \times 64 \times 64$ since the qualitative results for different lattice sizes do not show significant differences.

Results

Fig. 1 illustrates snapshots of the histone H2AX for different representative temperatures (in reduced units) in the range of $T^* = 0.014 - 0.025$. Some of the general conformational characteristics such as globular structure formation (global aggregation of intra-chain residues, $T^* = 0.014$), local segregation of selective residues ($T^* = 0.016$), large-scale stretching ($T^* = 0.020$), onset of randomization ($T^* = 0.022$) and thermal mixing ($T^* = 0.025$) are already apparent in the "snapshot" configurations. The interplay between the cooperative and competing interactions among the residues and the temperature constrained by the peptide bonds leads to a rich ensemble of protein structures. While a detailed analysis of such a structural ensemble in an in vivo system still remains an open challenge, our approach offers some insights into the overall structural pattern changes that are so far inaccessible to experiment. Moreover, simulation studies (such as ours) preceding experimental tests may help e.g. in assessing the applicability of proteins for the design and fabrication of biomolecular devices in the bionanotechnology.

Fig. 2 and Fig. 3 show the energy and mobility profile of each residue. The energy of a residue is its interaction energy with
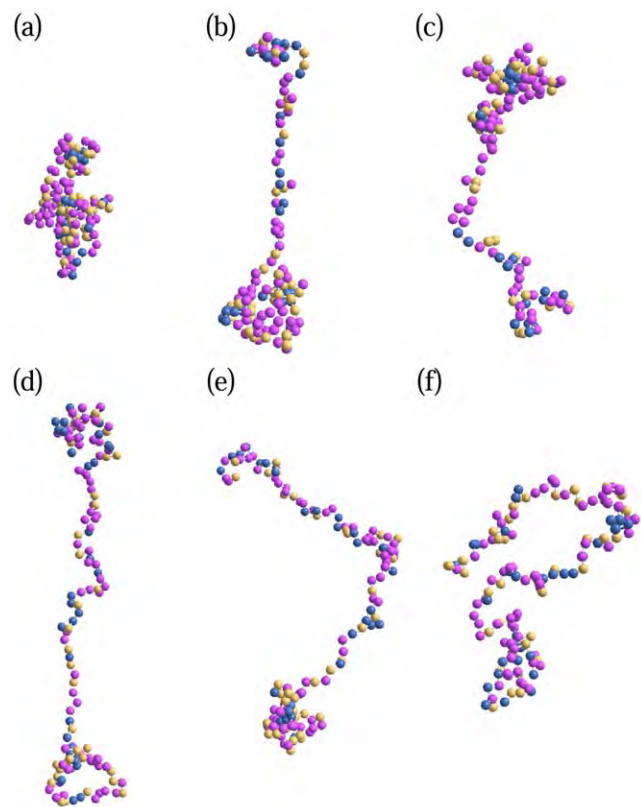

Figure 1. Snap shots of the histone H2AX at (a) $T^* = 0.014$, (b) $T^* = 0.015$, (c) $T^* = 0.016$, (d) $T^* = 0.020$, (e) $T^* = 0.022$ and (f) $T^* = 0.025$. Hydrophobic residues are shown in pink, polar residues in gold and electrostatic ones in blue. We find globular configurations at low temperature ($T^* = 0.014$), combination of segmental globules and elongated chains ($T^* = 0.015, 0.016$), predominantly elongated chains ($T^* = 0.020$), as well as universal SAW conformations at high temperature ($T^* \geq 0.022$).
doi:10.1371/journal.pone.0032075.g001

neighboring residues within the range of interaction. The mobility of a residue is defined by the number of successful moves per unit MC time step. Note the contrasts in profiles at relatively low ($T^* = 0.014, 0.015$) and high ($T^* = 0.025$) temperature. Residues along the histone backbone appear to possess an isotropic distribution of (almost in equal number) attractive (cohesive, negative) and repulsive (positive) energy. The magnitude of the repulsive and the attractive energy and their differences in consecutive segments increases with temperature which is manifested in the segmental configuration as well as in the global (coil-to-globule) structure of the protein.

The mobility profile of the residues follows the energy profile relatively closely where the lower segmental energy differences translate into lower segmental mobility. At low temperatures ($T^* = 0.014, 0.015$), residues with the lowest mobility consist of 36R, 37K, 57E, 62E, 65E, 72R, 73D, 75K, 76K, 78R, 89R, 90N, 91D, 92E, 93E, 95N, 96K, 119K, 120K, 134K, 135K, 141Q and 142E. Nearly all the electrostatic residues (D, E, K, R) along with a few polar groups (Q,N) act as anchor/seed for segmental aggregation. Note that the pair interaction potentials of these residues have the largest well depth [19]. Most of these residues become more mobile when raising the temperature (i.e. see the segments 91D, 92E, 93E).

It should be pointed out that some residues (e.g. 42E and 57E) have surprisingly low mobility despite their positive energy while
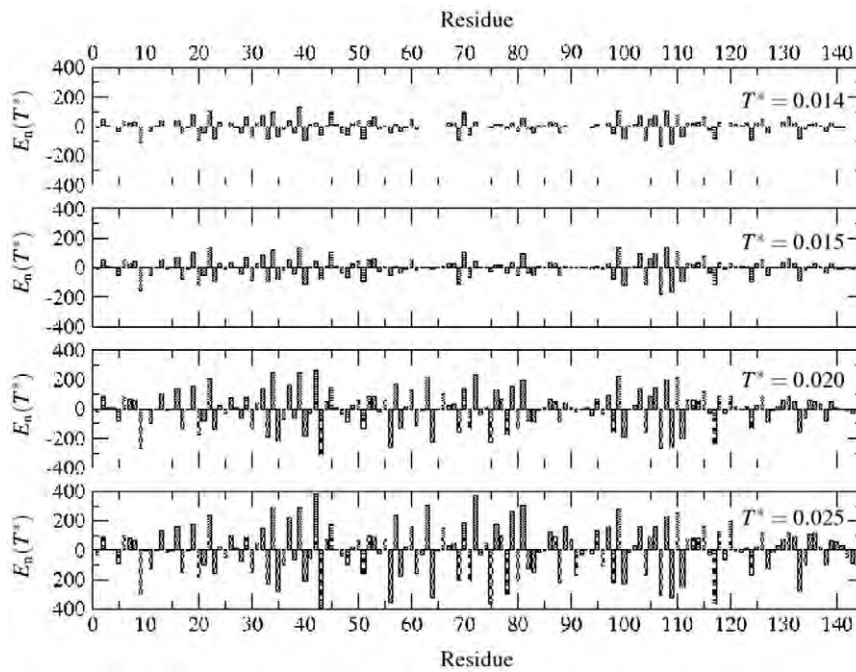
Figure 2. Energy E_n of each residue of histone H2AX at normalized temperatures $T^* = 0.014, 0.015, 0.020,$ **and** $0.025.$ The energy of a residue is its interaction energy with neighboring residues within the range of interaction. The magnitude of the repulsive and the attractive energy and their differences in consecutive segments increases with temperature which is manifested in the global (coil-to-globule) structure of the protein. At low temperatures ($T^* = 0.014, 0.015$), residues with the lowest mobility consist of 36R, 37K, 57E, 62E, 65E, 72R, 73D, 75K, 76K, 78R, 89R, 90N, 91D, 92E, 93E, 95N, 96K, 119K, 120K, 134K, 135K, 141Q and 142E. Nearly all the electrostatic residues (D, E, K, R) along with a few polar groups (Q,N) act as anchor/seed for segmental aggregation.
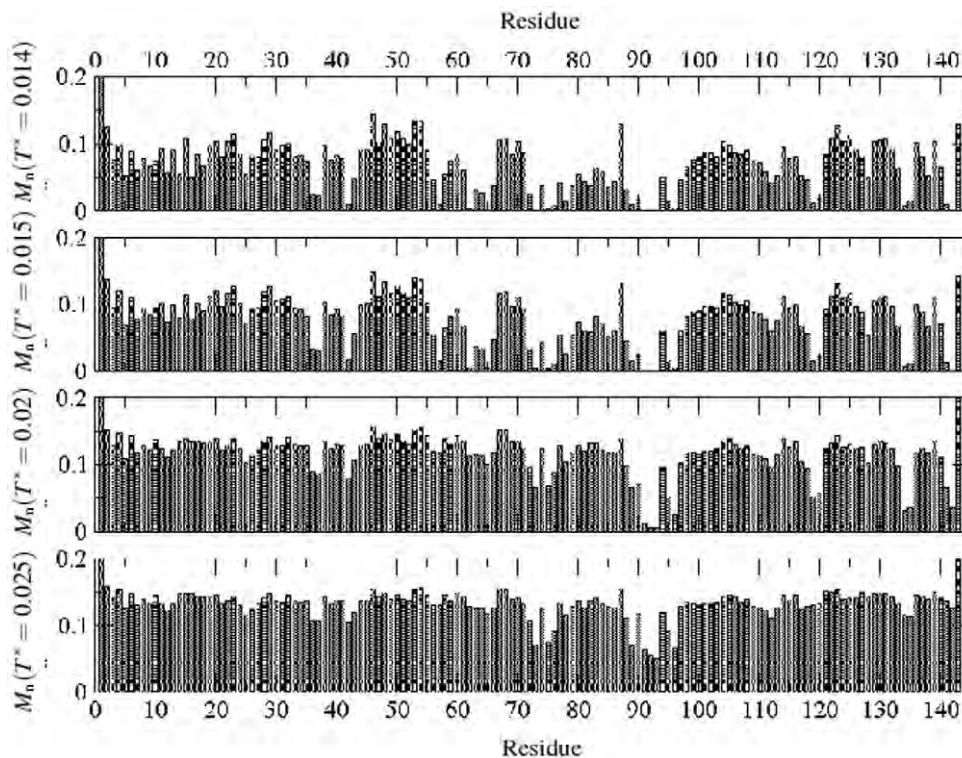doi:10.1371/journal.pone.0032075.g002

Figure 3. Mobility M_n of each residue (number of successful moves per unit time step) of the histone H2AX at temperatures $T^* = 0.014, 0.015, 0.020,$ **and** $0.025.$ The mobility of a residue is defined by the number of successful moves per unit MC time step. With exceptions the mobility profile of the residues follows the energy profile relatively closely where the lower segmental energy differences translate into lower segmental mobility. Interestingly, some residues, 42E and 57E, show a low mobility despite their positive energy.
doi:10.1371/journal.pone.0032075.g003

others with low energy have a high mobility index. In fact, conformational energy of the amino acids sequence (the interaction energy) does not determine the local structure and mobility alone. Physical (covalent bonding) or topological (trapping) constraints also play an important role in the cooperative response.

With respect to biotechnological applications it is interesting to study how the entire range of temperatures affects a protein's size and shape. The temporal variation of the radius of gyration R_{gyr} shows that it has reached its equilibrium at all temperatures except for the lowest one ($T^* = 0.014$) where the relaxation is too slow. Equilibration implies that the protein chain has explored a sufficient amount of conformations in structural phase space. The average value of the equilibrium radius of gyration can be evaluated from the asymptotic data sets at each temperature. R_{gyr} shows a non-monotonic dependence on temperature as can be seen in Fig. 4 with a maximum at a characteristic temperature T_c^* which is a "unique" property of the studied biomaterial. The radius of gyration increases on increasing the temperature ($T^* \approx 0.014 - 0.019$) from the low end until around $T_c^* \approx 0.019$ followed by a linear decay ($T^* \approx 0.019 - 0.023$) before reaching its saturation at high temperature $T^* \geq 0.023$.

We examine how the competition between residue-residue interactions and thermal fluctuations leads to the observed non-monotonic temperature dependence of the radius of gyration. The attractive inter-residue interaction induces self-assembly of the protein segments towards a global globular structure ($T^* \rightarrow 0.014$)

as the residues undergo their stochastic motion. As discussed above the highly interacting electrostatic residues act as an anchor collecting even those tethered residues that are repulsive. Thus, cooling down the protein leads to a conformational collapse into its globular conformation. The self-organized protein structure in its globular conformation begins to break on raising the temperature even by a small amount ($T^* \approx 0.016$) as the constitutive residues dissociate while some local assembly (held together by non-covalent interactions) still persists. The local assembly de-segregates on further increasing the temperature ($T^* = 0.017$) which stretches the corresponding segments resulting in a larger radius of gyration. Stretching of the protein continues until the characteristic temperature ($T_c^* \approx 0.019$) is reached beyond which the protein chain begins to contract. The chain segments fluctuate introducing randomness into a relatively stable elongated structure as the onset of thermal fluctuations sets in. The process can be described as a "thermal-driven contraction" emerging due to the cooperative effect of segmental interaction (looping) and conformational entropy. Note that the protein H2AX expands (positive) on heating in the low temperature regime ($T^* \approx 0.014 - 0.019$) and contracts (negative) at the higher temperatures ($T^* \approx 0.019 - 0.023$).

The crossover from a positive thermal response in low temperature to a negative thermal response in high temperature regimes with a well-defined transition temperature ($T_c^* \approx 0.019$) appears to be a specific characteristic of the H2AX protein structure. The results of another protein from histone family H3.1 provides additional support to our finding that histone H2AX
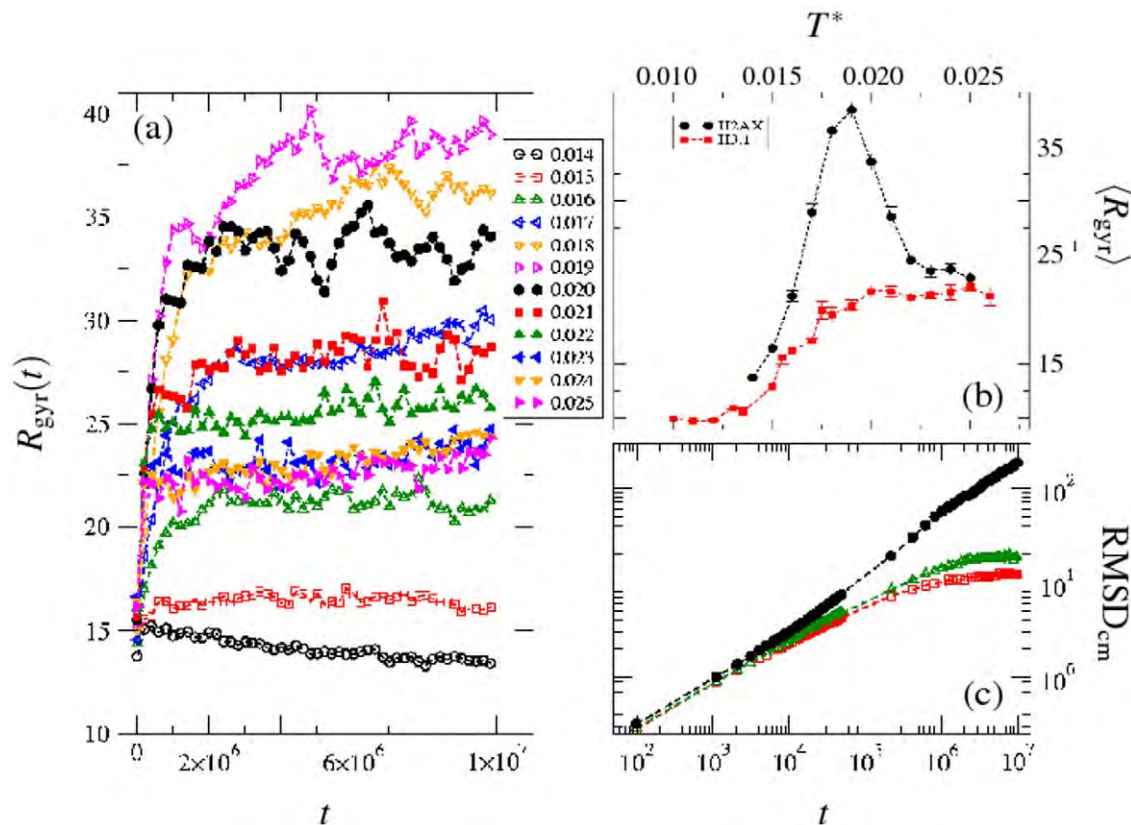

Figure 4. (a) Dependence of the radius of gyration $R_{gyr}(t)$ on MC time steps t for various temperatures $T^* = 0.014 - 0.025$; (b) temperature-dependent behavior of the mean radius of gyration $\langle R_{gyr} \rangle$ of the protein H2AX compared to the protein H3.1 where the same interaction parameters and normalized simulation temperature have been applied; (c) root mean squared displacement of the protein's center of mass $RMSD_{cm}$ as a function of MC time steps t. The crossover from a positive thermal response of $\langle R_{gyr} \rangle$ at low temperature to a negative thermal response at high temperature with a well-defined transition temperature $T_c^* \approx 0.019$ appears to be a specific characteristic of the H2AX protein structure.
doi:10.1371/journal.pone.0032075.g004

shows a unique peak in its R_{gyr}-temperature profile. The same interaction parameters and normalized simulation temperature are applied in the study of H3.1, which is of comparable size and does not show any peak in its R_{gyr}-temperature profile but exhibits a continuous transition from coil-to-globule on reducing the temperature. The difference in thermal response of H2AX and H3.1 however leads us to believe that this is due to the specific sequence of amino acids which might be exploited in technological applications requiring a material with such a distinct temperature response. Eventually, the radius of gyration saturates on further increasing the temperature beyond $T^* \geq 0.023$, where the protein conforms to a thermal-driven random (coil) structure.

The size of the protein as measured by its radius of gyration R_{gyr} can be compared at a low temperature ($T^* \leq T_c^*$) in the positive thermal response regime and at a higher temperature ($T^* \geq T_c^*$) in the negative thermal response regime. One has to point out that, despite having the same magnitude of R_{gyr}, the structure of the protein at these temperatures is very different. In particular, as shown in Fig. 1 we observe a local segmental segregation at low temperature while random configuration at high temperature dominate.

Although the radius of gyration can provide insight into the spatial extension of the protein, the specific dynamics of local structures are difficult to quantify with this measure. Thus, we have analyzed the root mean squared displacement of the center of mass of the protein with as a function of time for the entire temperature range. Fig. 4 shows these results for representative temperatures. The protein continues to diffuse at high temperatures while its motion slows down on reducing the temperature showing sub-diffusive asymptotic dynamics. At very low temperatures, $T^* = 0.014 - 0.015$, the dynamics are too slow since the protein is localized into its globular conformation.

The question has to be raised whether there is another property such as the specific heat C_V for which T_c^* has a special significance. C_V is evaluated from the fluctuation in the energy $\langle \Delta E^2 \rangle / T^2$ and Fig. 5 shows that the specific heat C_V does not

show a peak characteristic per se. This can be understood by noting that the characteristic temperature T_c^* is related to the maximum thermal response in the spatial extension and not to the identification of a phase transition. However, we see that C_V decays rather fast with increasing temperature before reaching a weave-like saturation. The undershoot in C_V occurs around T_c^* followed by an overshoot ($T \sim 0.0212$) which could be seen as a minimum and maximum, respectively, in the temperature profile of C_V. The thermal phase transition in general is associated with the divergence of the thermal correlation length at the critical point. The relaxation of the protein here is not only controlled by the competition between residue-residue interactions and temperature (i.e. the thermal fluctuation) but also by the steric constraints imposed by the peptide bonds (entropic barrier which is hard to evaluate). Moreover, the size of the protein is too small to identify the second order phase transition generally identified in macroscopically large (i.e. infinitely long ideal) chain systems. Thus, the thermal response in C_V is expected to be different from that of the radius of gyration which is a result of both interaction (thermal) and entropic contributions and therefore more reliable as a direct measure.

The structure factor $S(q)$ provides the spatial scaling of the distribution of constitutive elements

$$S(q) = \langle \frac{1}{N} | \sum_{j=1}^{N} \exp(-i\mathbf{q} \cdot \mathbf{r}_j)|^2 \rangle_{|\mathbf{q}|} \qquad (2)$$

where $\mathbf{r}_j$ is the position vector of each residue and $|\mathbf{q}| = q = 2\pi\lambda$ is the average spherical wave vector of wave length λ. For the protein, the structure factor is useful in understanding the structural details over a range of length scales. Fig. 6 shows the variation of $S(q)$ with the wave vector on a log-log scale. Since the radius of gyration R_{gyr} is a measure of the residue spread, the distribution of residues in the range of $q = 0.17 - 0.50$ provides information about the global conformation of the protein. From
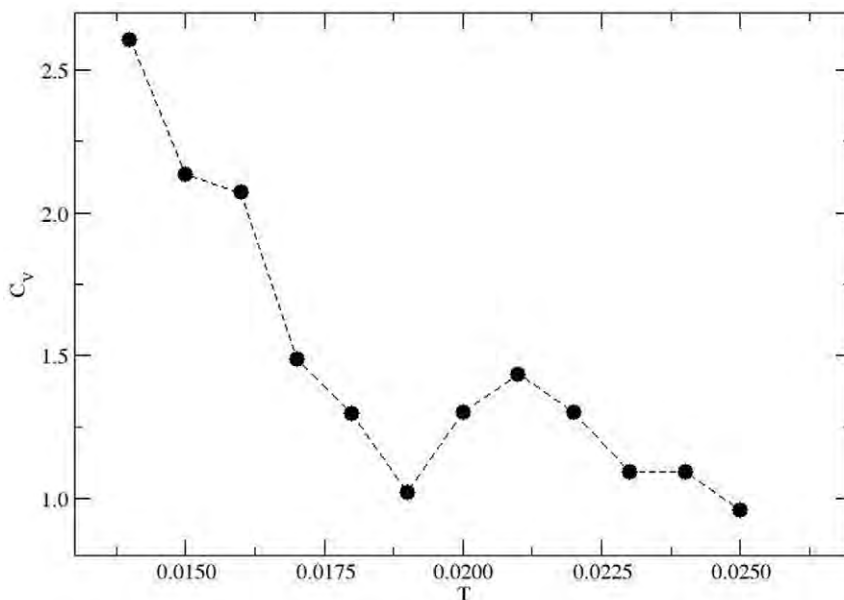

Figure 5. C_V is evaluated from the fluctuation in the energy $\langle \Delta E^2 \rangle / T^2$ and shows a minimum in C_V around T_c^*. In fact, the relaxation of the protein is not only controlled by the competition between residue-residue interactions and temperature but also by the steric constraints imposed by the peptide bonds.
doi:10.1371/journal.pone.0032075.g005

the power-law scaling of the structure factor with the wave vector, $S(q) \sim q^{-1/\nu}$, we can estimate the distribution of the protein segments $R_{\text{gyr}} \sim N^{\nu}$. We have estimated the slope of the power-laws in the appropriate range of the wave vector q in Fig. 6: $-1/\nu \approx 2.84$ at $T^{*\text{low}} = 0.014$ and $-1/\nu \approx 1.70$ at $T^{*\text{high}} = 0.025$ which provides $\nu^{\text{low}} = 0.35$ and $\nu^{\text{high}} = 0.59$, respectively. Thus, the protein has an effective dimension $D_{\text{eff}}^{\text{low}} \approx 1/\nu^{\text{low}} \approx 2.84$ (almost solid, a globular structure) at $T^* = 0.014$ and $D_{\text{eff}}^{\text{high}} \approx 1.70$ (a ramified, tenuous SAW structure) at $T^* = 0.025$. As shown in Fig. 6 a systematic change in the mass distribution of the protein is clearly seen on increasing the temperature ($T^* = 0.014 - 0.025$). On the lower spatial scale (higher q) there are minor modifications in the mass distribution at $T^* = 0.014$ while the protein segments appear like an ideal chain with $\nu \approx 1/2$ at $T^* = 0.025$.

Discussion

In summary, a coarse-grained protein model is used to study the structural dynamics of the histone H2AX which consists of 143 residues tethered in a bond fluctuating chain on a cubic lattice. Although the atomistic details of residues are ignored, their specificity is captured via a knowledge-based interaction matrix as well as a LJ pair potential for residue-residue interactions. Each residue executes its stochastic motion according to the Metropolis criterion.

We have analyzed a number of local and global physical quantities such as the energy and mobility of each residue as well as the root mean squared displacement of the protein's center of mass, its radius of gyration, and its structure factor. The impact of temperature on these quantities is investigated and might be exploited for the design of biomaterials.

Our approach allows for the identification of segmental characteristics such as active regions and anchoring sites of the protein. We find that the electrostatic residues (e.g. 72R, 73D, 91D, 92E, 93E, 134K, 135K, etc.) are critical in orchestrating the segmental conformation, their self-assembly and de-segregation from the low to the moderately high temperature regime ($T^* = 0.014 - 0.025$). These highly interacting residues at their specific positions in the protein sequence appear to determine specificity and multi-scale structures. Accordingly, we observe global globular configurations at low ($T^* = 0.014$), a combination of chains segments and smaller segmental globules at intermediate ($T^* = 0.016, 0.017$), and elongated structures at moderately high temperatures ($T^* = 0.019, 0.020$). As expected, the specificity of residues vanishes at high temperatures ($T^* \geq 0.023$) where the mobility of most residues becomes considerably high and comparable. In this thermal-driven structural regime, the residues become indistinguishable leading to a SAW chain conformation.

The radius of gyration of the protein shows a non-monotonic dependence on the temperature with a maximum at a characteristic temperature which is determined by the competition between inter-residue interactions and temperature. The protein H2AX expands (positive thermal response) on heating in the low temperature regime ($T^* = 0.014 - 0.019$) and contracts (negative thermal response) at higher temperatures ($T^* = 0.019 - 0.023$). The crossover from a positive to negative thermal response occurs at a well-defined transition temperature ($T_c^* \approx 0.019$) which may be a specific characteristic of the histone H2AX and particularly interesting for bioengineering purposes. The variation in the global conformation of the protein is explained in the framework of self-assembly at the local scale.

Based on the analysis of the structure factor $S(\mathbf{q})$, we find that the radius of gyration scales with its molecular weight N as $R_{\text{gyr}} \sim N^{\nu}$, where $\nu = 0.35$ and $\nu = 0.59$ at $T^* = 0.014$ and 0.025, respectively. The effective dimension of the protein is therefore $D_{\text{eff}} \approx 2.84$ (almost solid, a globular structure) at $T^* = 0.014$ and
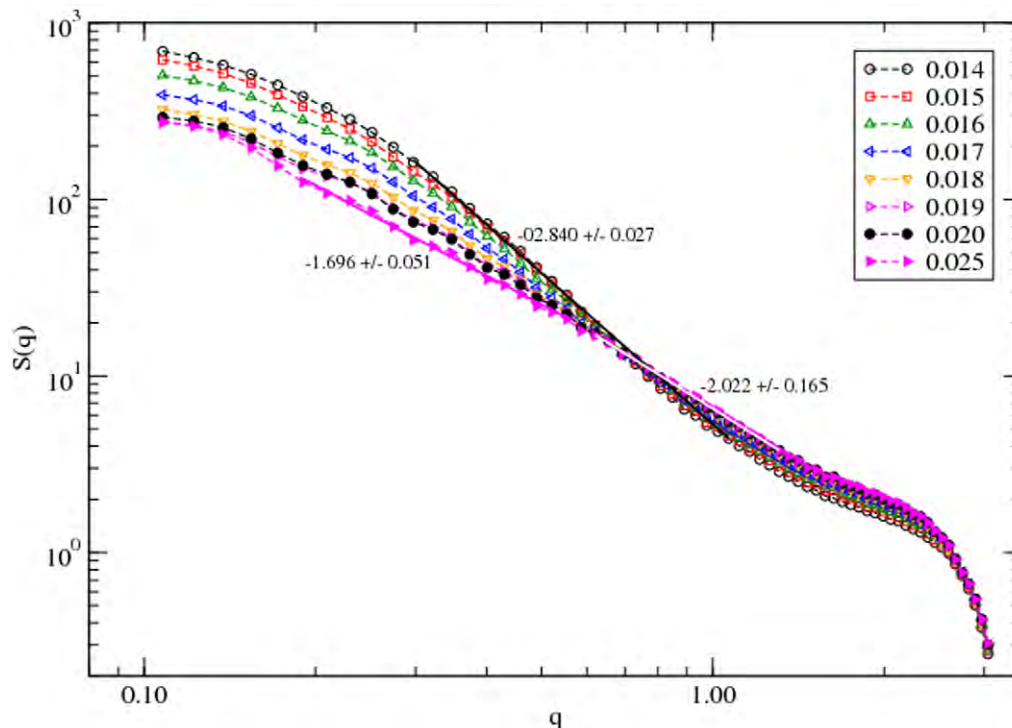

Figure 6. Structure factor $S(q)$ versus wave vector q. The effective dimension of the protein is $D_{\text{eff}} \approx 2.84$ (almost solid, a globular structure) at $T^* = 0.014$ and $D_{\text{eff}} \approx 1.70$ (a ramified, tenuous SAW structure) at $T^* = 0.025$.
doi:10.1371/journal.pone.0032075.g006

$D_{\text{eff}} \approx 1.70$ (a ramified, tenuous SAW structure) at $T^* = 0.025$. A systematic change in the mass distribution is clearly seen with an increase in temperature ($T^* = 0.014 - 0.025$).

Our coarse-grained protein model allows for a deeper understanding of local and global properties, which can so far not be gained by experimental testing. Besides the biological importance of proteins such as H2AX, we are able to provide a framework for analyzing potential candidates for the bioengineering of nano-materials. To this end, future experiments measuring

physical quantities such as the spatial extension (radius of gyration) as a function of temperature would allow for the calibration of the temperature scale.

Author Contributions

Conceived and designed the experiments: RBP MF. Performed the experiments: RBP MF. Analyzed the data: RBP. Contributed reagents/materials/analysis tools: RBP MF. Wrote the paper: RBP MF DWH BLF.

References

1. Renugopalakrishnan V, Garuňo-Juárez R, Narasimhan G, Verma CS, Wei X, et al. (2005) Rational design of thermally stable proteins: Relevance to bionanotechnology. J Nanosci Nanotech 5: 1759–1767.
2. Wang N, Mattis DM, Sundberg EJ, Schlievert PM, Kranz D (2010) A single, engineered protein therapeutic agent neutralizes exotoxins from both staphylococcus aureus and streptococcus pyogenes. Clinical and Vaccine Immunology 17: 1781–1789.
3. Kobsa S, Saltzman W (2008) Bioengineering approaches to controlled protein delivery. Pediatric Research 63: 513–519.
4. Asuri P, Bale SS, Karajanagi SS, Kan R (2006) The protein-nanomaterial interface. Curr Opinion Biotechnology 17: 562–568.
5. Kane RS, Stroock A (2007) Nanobiotechnology: Protein-nanomaterial interactions. Biotechnology Progress 23: 316–319.
6. Korkegian A, Black ME, Baker D, Stoddard BL (2005) Computational thermostabilization of an enzyme. Science 308: 857.
7. Liang HK, Huang CM, Ko MT, Hwang JK (2005) Amino acid coupling patterns in thermophilic proteins. Proteins 59: 58.
8. Das R, Gerstein M (2000) The stability of thermophilic proteins: a study based on comprehensive genome comparison. Functional and Integrative Genomics 1: 76.
9. Paull TT, Rogakou EP, Yamazaki V, Kirchgessner CU, Gellert M, et al. (2000) A critical role for histone h2ax in recruitment of repair factors to nuclear foci after dna damage. Curr Biol 10: 886–895.
10. Fernandez-Capitello O, Chen HT, Celeste A, Ward I, Romanienko PJ, et al. (2002) Dna damageinduced g2-m checkpoint activation by histone h2ax and 53bp1. Nature Cell Biol 4: 993–997.
11. Bewersdorf J, Bennett BT, Knight KL (2006) H2ax chromatin structures and their response to dna damage. PNAS 103: 18137–18142.
12. Ikura T, Tashiro S, Kakino A, Shima H, et al. (2007) Dna damage-dependent acetylation and ubiquitination of h2ax enhances chromatin dynamics. Mol And Cell Biol 27: 7028–7040.
13. Van den Burg B, de Kreij A, van der Veek P, Manfeld J, Venema G (1999) Characterization of a novel stable biocatalyst obtained by protein engineering. Biotech Appl Biochem 30: 35.
14. Graddis TJ, Remmele RL, McGrew JT (2002) Designing proteins that work using recombinant technologies. Curr Pharm Biotechnol 3: 285.
15. Somero GN (1995) Proteins and temperature. Annu Rev Physiol 57: 43–68.
16. Smerdon M, Isenberg I (1973) The effect of temperature on histone grk aggregation. Biochemical and Biophysical Research Communications 55: 1029–1034.
17. Eickholt J, Wang Z, Cheng J (2011) A conformation ensemble approach to protein residue-residue contact. BMC Struct Biol 11: 1–8.
18. Pandey RB, Farmer BL (2010) Globular structure of a human immunodeficiency virus-1 protease (1difa dimer) in an effective solvent medium by a monte carlo simulation. J Chem Phys 132: 125101–125105.
19. Miyazawa S, Jernigan RL (1985) Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. Macromolecules 18: 534–552.
20. Eby DM, Johnson GR, Farmer BL, Pandey RB (2011) Supramolecular assembly of a biomineralizing antimicrobial peptide. Phys Chem 13: 1123–1130.
21. Feng J, Pandey RB, Berry RJ, Farmer BL, Naik RR, et al. (2011) The mechanism of adsorption conforms to soft epitaxy observed for peptides on metal surfaces. Soft Matter 7: 2113–2120.
22. Pandey RB, Farmer BL (2008) Phys Rev E 77: 031902–031910.
23. Lyubartsev A, Laaksonen A (1995) Calculation of effective interaction potentials from radial distribution functions: A reverse monte carlo approach. Phys Rev E 52: 3730–37.
24. Izvekov S, Voth G (2005) J Phys Chem B 109: 2469–73.
25. Zhou J, Thorpe I, Izvekov S, Voth G (2007) Coarse-grained peptide modeling using a systematic multiscale approach. Biophys J 92: 4289–03.
26. Zhou J, Chen S, Jiang S (2003) Orientation of adsorbed antibodies on charged surfaces by computer simulation based on a united-residue model. Langmuir 19: 3472.
27. Marrink S, de Vries A, Mark A (2004) Coarse grained model for semi-quantitative lipid simulations. J Phys Chem B 108: 750–60.
28. Marrink S, Risselada H, Yefimov S, Tieleman D, de Vries A (2007) The martini force field: coarse grained model for biomolecular simulations. J Phys Chem B 111: 7812–24.
29. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM (1998) Dna double-stranded breaks induce histone h2ax phosphorylation on serine 139. J Biol Chem 273: 5858–5868.
30. Ayoub N, Jeyasekharan AD, Bernal JA, Venkitaraman AR (2008) Hp1-beta mobilization promotes chromatin changes that initiate the dna damage response. Nature 453: 682–686.
31. Kongruttanachok N, Phuangphairoj C, Ponyeam W, Mutirangura A (2008) Replication independent dna double-strand break retention may prevent genomic instability. Science Asia 34: 253–257.
32. Tetievsky A, Horowitz M (2010) Posttranslational modifications in histones underlie heat acclimation-mediated cytoprotective memory. J Appl Physiol 109: 1552–1561.
33. Knowledge-based potential server for proteins at coarse grained level. Available: http://gor.bb.iastate.edu/potential/. Accessed 2010 Nov.
34. Carmesin I, Kremer K (1988) The Bond Fluctuation Method: A New Effective Algorithm for the Dynamics of Polymers in All Spatial Dimensions. Macromolecules 21: 2819–2823.
35. Binder K, Heermann DW (2010) Monte Carlo Simulation in Statistical Physics: An Introduction. Berlin: Springer Verlag, 5 edition.
36. Bjursell J, Pandey RB (2004) Relaxation to native conformation of a bond-uctuating protein chain with hydrophobic and polar nodes. Phys Rev E 70: 052904.
37. Chen CM, Chen CC (2003) Computer simulations of membrane protein folding: structure and dynamics. Biophys Journal 84: 1902.
38. Miyazawa S, Jernigan RL (1996) Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. J Mol Biol 256: 623–644.
39. Behringer H, Degenhard A, Schmid F (2006) Coarse-grained lattice model for molecular recognition. Phys Rev Lett 97: 128101–128104.
40. Skepo M, Linse P, Arnebrant T (2006) Coarse-grained modeling of proline rich protein 1 (prp-1) in bulk solution and adsorbed to a negatively charged surface. J Phys Chem 110: 12141–12148.
41. Zhou Y, Karplus M (1999) Interpreting the folding kinetics of helical proteins. Nature 400: 400–403.
42. Shen M, Freed KF (2002) All-atom fast protein folding simulations: the villin headpiece. Proteins 49: 439–445.
43. Wei Y, Nadler W, Hansmann UHE (2008) Backbone and side-chain ordering in a small protein. J Chem Phys 128: 025110–025115.
44. Chapagain PP, Gertsman B (2006) Removal of kinetic traps and enhanced protein folding by strategic substitution of amino acids in a model alpha-helical hairpin peptide. Biopolym 81: 167–178.
45. Wang Y, Feng S, Voth G (2009) Transferable coarse-grained models for ionic liquids. Chem Theory Comput 5: 1091–98.
46. Carbone P, Varzaneh H, Chen X, Müller-Plathe F (2008) Transferability of coarse-grained force fields: The polymer case. J Chem Phys 128: 064904.
47. Gerstman BS, Chapagain PP (2005) Self-organization in protein folding and the hydrophobic interaction. J Chem Phys 123: 054901–6.